

---

## UNIT 6 MEASUREMENT OF VARIABLES

---

### Structure

- 6.0 Objectives
- 6.1 Introduction
- 6.2 Types of Variables
- 6.3 Measurement of Qualitative Data
- 6.4 Census versus Sample Survey
- 6.5 Sampling Procedure
- 6.6 Types of Sampling
- 6.7 Summary
- 6.8 Answers to Self Check Exercises
- 6.9 Keywords
- 6.10 References and Further Reading

---

### 6.0 OBJECTIVES

---

After going through this Unit, you will be able to:

- define a variable;
- distinguish between various types of variables;
- measure ordinal variables through scaling techniques;
- distinguish between census and sample survey;
- explain the steps involved in carrying out a sample survey; and
- distinguish between various types of sampling.

---

### 6.1 INTRODUCTION

---

During the course of research you come across situations where you have to measure various characteristics. These characteristics could be of various types, viz., age, height or income level of visitors to a library; educational qualifications, social status or reading habits of a person; gender, religion, or area of interest of a library user. Note that all these characteristics are not similar from the point of measurement. While age, height or income can be measured in quantitative terms (in number of years, in centimetres, in rupees) religion or gender can be put to certain categories only. In this Unit we discuss the issue of measurement of qualitative variables, particularly the scaling techniques.

Another issue discussed in the Unit pertains to collection of data on the basis of sample survey. Very often it is not possible to survey all the units bearing the characteristic under study. The constraints could be inadequate funds, time limit, and manpower. In such situations we survey only a subset of the population, called sample. We discuss various concepts associated with sampling procedure.

We begin with the types of variables and their measurement.

## 6.2 TYPES OF VARIABLES

Let us begin with the concept of a variable. It is a characteristic of the *sample* or the *population* that we intend to measure. Thus age of the reader is a variable, so is gender, educational level or mother tongue. As we have mentioned earlier all variables are not similar.

Variables can be of two types: qualitative and quantitative. Qualitative variable is one that cannot be expressed in numerical terms. For example, marital status is a qualitative variable. Here we can have two categories: married and single. Of course, if you want a more detailed categorisation you can further divide single in to widow/widower, divorcee and never married. Similarly, gender (male or female), mother tongue (Hindi, Bengali, Oriya, Tamil, Urdu, etc.), subject categories (economics, history, physics, medicine, etc.), religion (Buddhism, Christianity, Hinduism, Islam, etc.) are examples of qualitative variables. Here we study an attribute or quality that cannot be quantified, but can be divided into various categories. Moreover, we cannot say that one category is higher or greater than another category. Such variables are also called nominal variables.

There is another type of qualitative variable where we can divide the observations into various categories and also say that one category is higher or greater than another category. An example could be the educational qualification of a visitor to a library. Here we can divide the visitors on the basis of their educational qualification into categories such as 'secondary', 'senior secondary', 'graduate' and 'post-graduate'. In this case, obviously, the category 'Senior Secondary' is higher than the category 'Secondary' in terms of number of years of schooling and expected mental maturity. In this case we arrange the categories in an ascending or descending order. This sort of variables are called ordinal variables.

In the case of nominal variables we cannot perform any mathematical operations (such as addition, subtraction, multiplication, division,) or logical operations (greater than, less than) across categories. We can simply count the number of observations in each category. In the case of ordinal variables we can say that one category is greater than another category. But we cannot quantify the difference between categories. For example, we cannot express numerically the difference between two categories (say secondary and senior secondary). Also we cannot say that the difference between two categories (say secondary and senior secondary) is the same as the difference between two other categories (say graduate and post-graduate).

A quantitative variable can be expressed in numerical terms. Hence it is also called numerical variable. Examples of numerical variable could be age, income, weight, height, distance travelled, etc. This category of variables can be subjected to various mathematical and logical operations. Thus we can express the monthly income of a librarian in Rupees and also say by what percentage it exceeds the salary of a library assistant.

Numerical variables can be of two types: discrete and continuous. Discrete variable is one where the observations assume values in complete numbers. For example, the number of children in family can only be whole numbers; it cannot be fractions. On the other hand, continuous variables can assume any value in an interval. For example, weight of a person can be measured to any precision and thus can take any value in between two points.

Let us distinguish between variable and data. We obtain data by measuring a variable (qualitative or quantitative) on certain individuals or units. For example, suppose we measure the height of 50 employees in a library. Here height is the variable and 50

**Quantitative Research.** These 50 numerical values that we obtain are our data. Thus we have discrete data or continuous data depending upon whether the variable is discrete or continuous.

Similarly there are primary data and secondary data. Primary data refers to data collected by the researcher by undertaking a field survey. On the other hand, secondary data refers to collection of data from published sources, e.g., census, budget, handbooks, etc. Thus when you undertake a field survey, collect data, analyse the results and present it in some forum, it is primary data. But when I use that data for further analysis it becomes secondary data for me.

---

### 6.3 MEASUREMENT OF QUALITATIVE DATA

---

The measurement of quantitative data does not pose problems as these are expressed in numerical terms. Measurement of qualitative data, however, is a complex issue and needs to be discussed further. There are two types of variables that are usually measured: i) social behaviour and personality, and ii) cultural and social environment. The purpose is to study socioeconomic status and its impact on various issues, attitude towards a particular event or behaviour, participation in certain activity, etc.

In order to measure these types of issues no readymade scale of measurement is available. Secondly, there is a lot of ambiguity and inter-personal variability in the definition of these qualities. However, attempts have been made to construct 'attitude scales' to measure these qualities. These scales are in the form of equal-appearing intervals. Such scales are widely used in sociology, psychology, education and administration apart from other branches of study.

In designing such scales we should keep one thing in mind that it should be 'valid' and 'reliable'. In other words, i) it should measure what we intend to measure, and ii) it should yield the same consistent results when applied under the same conditions.

In designing attitude scales we form a number of statements and ask the respondents to react to these statements. The statements should be brief, unambiguous and relevant. It should be expressed in such a form that it could be endorsed or rejected in terms of definitely expressed attitude.

A widely used scale is 'Likert Scale' which is also referred to as 'technique of summated ratings'. The basic steps in construction of Likert scale are given below:

A series of propositions representing attitudes are compiled. The attitudes of persons towards these propositions could range from extremely negative to extremely positive.

The statements or propositions express values rather than facts. Each statement indicates the position of a person towards the issue concerned. For example, a statement could be 'there should be separate reading rooms for boys and girls in a library'. The response to such a statement could range from 'strongly agree' to 'strongly disagree' depending upon the attitude of the respondent towards the issue.

Each statement is so formulated that the response of persons can be given in any of the five terms such as i) *strongly approve*, ii) *approve*, iii) *undecided*, iv) *disapprove*, v) *strongly disapprove*. Many times expressions such as i) *strongly agree*, ii) *agree*, iii) *cannot say*, iv) *disagree*, v) *strongly disagree* are also used. Although a five-point continuum is common, there can be three-, four-, or six-point continuum also. For example, a seven-point continuum could be i) *always*, ii) *almost always*, iii) *frequently*, iv) *occasionally*, v) *rarely*, vi) *almost never*, vii) *never*.

Weights are assigned to each response for a statement. For example, Measures of Variables continuum the weights that we attach to each response could be 1, 2, 3, 4, and 5. These weights indicate the intensity of the attitudes of a respondent.

When we have a number of statements, we can arrange the statements according to their relative intensity. Also groups of individuals can be arranged in a rank order on the basis of scale scores. Moreover, the scale scores can be used to predict any outside variable.

### Self Check Exercise

- 1) State the type of variable (nominal, ordinal and numerical) for the following:
  - a) tribes in India
  - b) height of children
  - c) number of books
  - d) accession number of books
  - e) subject codes
- 2) With an example explain the use of Likert scale.

**Note:** i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of the Unit.

.....  
.....  
.....  
.....

## 6.4 CENSUS VERSUS SAMPLE SURVEY

There are quite a few methods for collection of data that we will discuss in detail in Block 3 (questionnaire method, personal interview, participatory observation, etc.). An important issue is whether we should collect data on all units present or only a part of it. For example, if we wanted to know the reading habits of economics students of Sambalpur University and for that purpose we designed a questionnaire. We found that there are 800 economics students and it would not be possible to survey all the students due to time and money constraints. Thus, we decided to administer the questionnaire to 100 students only.

In formal statistical language we distinguish between 'population' and 'sample'. Since our objective is to study the reading habits of economics students in Sambalpur University, all 800 economics students qualify to be studied by us and thus constitute our population. If we decide not to collect data on all students but study only 100 students, these selected 100 students constitute our sample.

Thus, population is a collection of individuals or objects having the desired characteristics we are interested in a sample which is a part of the population. Obviously we can draw more than one sample from the same population. If you go to Sambalpur University on a different day and select 100 students your sample may be different from what another person had selected.

**Collection of Information** on all units of a population is called census. On the other hand, collection of information on a sample is called sample survey. The process of selection of a sample is called 'sampling'. The advantages of sample survey are:

- Sample survey is less expensive than a census.
- Sample survey requires lesser time and manpower than a census.
- Sample survey can be monitored closely and more accurate information can be collected than in census.
- Many times we do not require information on all units. For finding out the effectiveness of a new medicine we do not have to test it on all patients; only a representative sample will do.

In sampling we are surveying only a part of the population. Remember that we do not know the exact population characteristics, as we do not survey the population. But there are scientific methods by which we can estimate a population characteristic (called parameter) on the basis of sample characteristic (called statistic). This topic will be covered in Unit 17 while dealing with estimation and hypothesis testing.

There are two types of errors in a survey: sampling error and non-sampling error. Sampling error is due to the fact that only a part of the population is being surveyed. It is the difference between the *parameter* and the *statistic*. It can be reduced by adoption of scientific sampling procedure, specifically probability sampling so that a random sample is selected. Non-sampling error is due to errors in measurement, non-response by the selected units, wrong recording of data, and personal bias of the investigators. Thus, non-sampling error is present in both census and sample survey. In a census, however, many investigators are involved and large number of units are to be surveyed. So non-sampling error could be very high. On the other hand, in a sample survey greater care can be taken in collection and recording of data since lesser number of units are surveyed. Thus non-sampling error can be minimised in a sample survey. The total error is the sum of sampling and non-sampling errors. It may so happen that the non-sampling error in a census is more than the sampling and non-sampling errors in a sample survey! In such cases sample survey would give better results than census.

### Self-Check Exercises

- 3) What are the advantages of sampling over census?
- 4) Distinguish between sampling error and non-sampling error.

**Note:** i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of the Unit.

.....

.....

.....

.....

---

## 6.5 SAMPLING PROCEDURE

---

While carrying out a sample survey there are certain steps that we should follow. These are as follows:

**Measurement of Variables**  
**Specification of objectives:** The objectives of carrying out sampling is the foremost step in sampling procedure. Because all other steps will follow from this objective.

**Definition of population:** In this step we define the units that should be included in the population. Many times there are certain border cases where proper definition is important. For example, if you want to survey the library personnel, it has to be decided whether you should include part-time employees in the population or not.

**Preparation of 'sampling frame':** Once you have defined the units to be included in the population, the next step is to prepare a list of the units from which the sample is to be drawn. Many times problems come up because the sources from which you want to prepare the sampling frame may be incomplete or obsolete.

**Identification of sampling procedure:** Sampling procedure refers to the method of selecting the sample. There are quite a few sampling procedures available. We should select a method that: i) gives us a representative sample, ii) is feasible to carry out keeping in view our constraints, and iii) is cost effective. There are broadly two types of sampling: probability sampling and non-probability sampling. We will discuss these types in the next Section.

**Determination of sample size:** The next step is the determination of sample size. The factors that influence sample size are i) population size, ii) variance of units, iii) desired precision level, iv) response rate, and v) availability of resources. The sample size should be relatively larger if population size is larger, variability among units is higher, more precise results are required, and lesser is the response rate. However, we should settle for smaller sample size if there are constraints such as availability of funds or manpower or time. Sample is considered to be large if its size is greater than 30. If sample size is less than or equal to 30 it is considered as small sample. The procedures of drawing inferences (See Unit 17) for large and small samples are different.

**Selection of sampling units:** Once you have decided on the sampling procedure and the size of the sample, the next step is draw the sampling units (that is, units included in the sample) from the sampling frame.

---

## 6.6 TYPES OF SAMPLING

---

As mentioned earlier there are two types of sampling: probability and non-probability. In probability sampling some element of randomness is involved in selection of units, so that personal judgement or bias is not there. Here we have simple random sampling, systematic random sampling, stratified random sampling and multistage sampling. Remember that the word 'random' here does not mean haphazard or unsystematic, rather it implies lack of bias or control in selection. In non-probability sampling we have judgement sampling, quota sampling, and snowball sampling.

**Simple Random Sampling (SRS):** It is the basic sampling procedure where each unit in the population gets an equal chance of being included in the sample. There are two commonly used methods to draw a simple random sample, viz., i) lottery method, and ii) random numbers selection method. In both the methods we assign a unique number to each unit in the sampling frame. In lottery method we mix up the numbers very well and draw the numbers one by one. In random number selection method we refer to 'random number tables (RNT)' available from various sources (including the Internet) and select the units which are there in the RNT. Simple random sampling should be used with a homogeneous population. That is, all the units in a population possess the same attributes that the researcher is interested in measuring. The characteristics of

**Tools for Research** include age, sex, income, social status, geographical region, etc. If the population is not homogeneous then we need a larger sample size to obtain a representative sample. But we can divide the population into various homogeneous strata and obtain a stratified random sample. A limitation of the simple random sampling is that if the population size is too large, selection of the sampling units is time consuming. These days, of course, computer has been very useful in selection of random numbers and reduced time spent in selection process.

**Systematic Random Sampling:** In this case we select the units in a fixed interval. For example, in a library you have to check the quality of maintenance of books. For this purpose you decide take physical verification of 100 books out of the 5000 books available in the library. Here population size is 5000 and sample size is 100. In systematic random sampling procedure you take a random starting point and then select the 50<sup>th</sup> subsequent unit. Thus if 12<sup>th</sup> book is the starting point, then you go on checking books located at the positions: 62<sup>nd</sup>, 112<sup>th</sup>, 162<sup>nd</sup>, ..... and so on. This procedure is useful when preparation of sampling frame is difficult or not possible.

**Stratified Random Sampling:** This procedure is practiced when the population is not homogeneous but can be divided into various homogeneous groups (called 'strata'). Here we select sub-samples from each stratum and add them together to obtain our desired sample. Therefore, the stratified random sampling procedure is a better representative of the population and sampled units reflect the population characteristics more accurately. The steps we should follow are:

- Divide the population into strata based on some characteristic chosen by you (example, Rural/Urban, Male/Female, etc.)
- The number of units taken from each stratum is proportional of the relative size of the stratum and standard deviation of the characteristic within the stratum. If stratum size is then a larger sub-sample should be taken. Similarly if you find that variability among units is more in a stratum than other strata, then you should take a larger sub-sample from that.
- Choose the sub-sample from each stratum using simple random sampling.

**Multistage Sampling:** As the name suggests, sampling procedure in this case is divided into two or more stages. Let us take an example of a two-stage sampling. Suppose you have to survey the aptitude of college students, who can be categorised into three streams: arts, science and commerce. Here in the first stage you divide colleges according to geographical regions and you select specified number of colleges from each region on the basis of random sampling. In the second stage you select students by a random sampling method from the colleges selected at the first stage only, not from all colleges. Let us try to explain the difference between two-stage sampling and stratified sampling. In stratified random sampling you divided the population (all college students) into there strata: arts, science and commerce. From each stratum you selected a sub-sample by simple random sampling procedure. The selected students can be from any college and you cannot rule out a visit to any college. In two-stage sampling you are excluding certain colleges at the first stage itself, which may prove to be economical.

**Judgement Sampling:** It is a non-probability sampling procedure. It is also called purposive sampling, where the researcher selects the sample based on his/her judgment. The researcher believes that the selected sample elements are representative of the population. For example, the calculation of consumer price index is based on judgment sampling. Here the sample consists of a basket of consumer items and other goods and services which are expected to reflect a representative sample. The prices of these items are collected from selected cities that are viewed as typical cities with

demographic profiles matching the national profile. The advantage of judgment sampling is that it is low cost, convenient and quick. The disadvantage is that it does not allow direct generalisations to population. The quality of the sample depends upon the judgment of the researcher.

**Quota Sampling:** In this procedure the population is divided into groups based on some characteristics such as gender, age, education, religion, income group, etc. A quota of units from each group is determined. The quota may be either proportional or non-proportional to the size of the group in the population. Do not confuse the quota sampling with stratified random sampling discussed earlier. In stratified random sampling you select random samples from each stratum or group whereas in quota sampling the researcher/interviewer has a fixed quota and selection of units is based on judgement. Quota sampling has the advantage that cost and time involved in selection of units is reduced considerably. However, it has many disadvantages also. In quota sampling, the samples are selected according to the convenience of the investigator. Therefore, the selected sample may be biased.

**Snowball Sampling:** In this procedure you begin by identifying someone who meets the criteria for inclusion in your study. You then ask them to recommend others who they may know who also meets the criteria. Although this method would hardly lead to representative samples, there are times when it may be the best method available. Snowball sampling is especially useful when you are trying to reach population that are inaccessible or hard to find. For example, if you are studying the homeless or people suffering from a particular disease you are may not be able to find good lists respondents. However, if you identify one or two, you may find that they know others in a similar condition.

### Self Check Exercise

- 5) What are the important steps in carrying out a sample survey?
- 6) Explain the procedure of drawing a stratified random sample.
- 7) Explain the following concepts:
  - a) Systematic random sampling
  - b) Parameter and statistic
  - c) Multistage sampling

**Note:** i) Write your answers in the space given below.  
ii) Check your answers with the answers given at the end of the Unit.

.....  
.....  
.....  
.....

---

## 6.7 SUMMARY

---

A variable is a characteristic that we are interested in analysing. It can be nominal, ordinal or numerical. Numerical variable can be discrete or continuous. While numerical variable is amenable to mathematical and logical operations nominal and ordinal variables are not. In certain cases it is difficult to measure ordinal variables because we do not

~~Two standards of measurement.~~ In such cases we need to construct a measurement scale. In this we discussed one such scaling techniques, that is summated rating.

Another issue that we covered in this Unit relates to sampling techniques. Because of inadequate resources or infeasibility we often resort to sampling instead of census of all the units in the population. There are two broad types of sampling procedures: probability sampling and non-probability sampling. In probability sampling we have simple random sampling, systematic random sampling, stratified random sampling and multistage sampling. On the other hand, non-probability sampling procedures are judgment sampling, quota sampling and snowball sampling.

---

## 6.8 ANSWERS TO SELF CHECK EXERCISES

---

- 1) a) Nominal, b) Numerical, c) Numerical, d) Nominal, and e) Nominal
- 2) Example of the use of Likert Scale: Show your agreement with the following:  
Libraries will ever remain important for the advancement of the society  
Strongly Agree; Agree; Undecided; Disagree; Strongly Disagree
- 3) Sampling is advantageous over statistic in that it saves cost and time to be devoted to the survey. It involves fewer personnel to be deployed and also results in more precision in the results.
- 4) Sampling error is due to the faulty sample selected. It may be due to non- probability sampling techniques adopted. Non- sampling errors are due to the errors in data measurement or analysis methods.
- 5) The steps to be followed in carrying out a sample survey are:  
specification of objective, preparation of sampling frame, identification of sampling procedure, determination of sample size, and selection of sampling units.
- 6) The procedure of drawing a stratified random sample is:
  - 1) Divide the population into strata based on some characteristic chosen by you (example, Post graduate/ Under graduate, Male/Female, etc.)
  - 1) Decide the number of units to be taken from each stratum proportional to the relative size of the stratum and standard deviation of the characteristic within the stratum. If stratum size is then a large a sub-sample should be taken. Similarly if you find that variability among units is more in a stratum than other strata, then you should take a larger sub-sample from that.
  - 1) Choose the sub-sample from each stratum using simple random sampling.
- 7) a) Systematic random sampling is a type of random sampling where the bias is minimised. Here the random sample is selected in a systematic way, e.g., in case a sample of 100 is to be selected from a population of 1000, and then we may select the first member of the population and every subsequent 20<sup>th</sup> member.
  - b) Parameter is a summary value of the population while statistic is that of the sample.
  - c) Multistage sampling is sampling done in two or more stages. In case we have to survey reading habits of users in public libraries, we may first take a random sample on the basis of geographical regions and then further take a sample on the basis of age groups. This is an example of multistage sampling.

## 6.9 KEYWORDS

- Convenience Sampling** : It refers to the method of obtaining a sample that is most conveniently available to the researcher.
- Judgement Sampling** : In this sampling procedure the selection of sample is based on the researcher's judgement about some appropriate characteristics required of the sample units.
- Multistage Sampling** : The sample selection is done in a number of stages.
- Parameter** : It is a measure of some characteristic of the population.
- Population** : It is the entire collection of units of a specified type in a given place and at a particular point of time.
- Quota Sampling** : In this sampling procedure the samples are selected on the basis of some parameters such as age, gender, geographical region, education, income, religion, etc.
- Sample** : It is a sub-set of the population. Therefore, it is a collection of some units from the population.
- Simple Random Sampling**: This is the basic sampling procedure where all units in the population have an equal chance of being included in the sample.
- Snowball Sampling** : It relies on referrals from initial sampling units to generate additional sampling units.
- Statistic** : It is a function of the values of the units that are included in the sample. The basic purpose of a statistic is to estimate some parameter.
- Stratified Random Sampling** : In this sampling procedure the population is divided into groups called strata. Subsequently sub-samples are selected from each stratum using a random sampling method.
- Systematic Sampling** : In this procedure the units are selected from the population at uniform interval (in time, order or space).

## 6.10 REFERENCES AND FURTHER READING

- Bhardwaj, R. S. (1999). *Business Statistics*. New Delhi: Excel Books.
- IGNOU Study Material (2005). *EEC-13: Elementary Statistical Methods and Survey Techniques*, Block 6.
- Kothari, C.R. (1985). *Research Methodology: Methods and Techniques*. New Delhi: Wiley Eastern.
- Young, P. V. (1988). *Scientific Social Surveys and Research*, Prentice Hall of India: New Delhi.